An Analytical Method to Determine Reference Orbitals for Localized Frontier Orbitals (LFOs) and Its Application to the Correlation Analysis between LFO Energies and Acidities of Conjugate Cations of Heterocyclic Bases and Monosubstituted Benzoic Acids

Yasuyuki Kurita*,† and Chiyozo Takayama‡

Agricultural Chemicals Research Laboratory, Sumitomo Chemical Co., Ltd., 4-2-1 Takatsukasa, Takarazuka, Hyogo 665, Japan

Received: March 12, 1997; In Final Form: May 23, 1997[⊗]

The variational method has been applied to the analytical determination of reference orbitals for localized frontier orbitals. Good linear correlations have been found between localized frontier orbital energies and acidities of conjugate cations of heterocyclic bases and monosubstituted benzoic acids.

Introduction

The frontier molecular orbital (FMO) theory is useful to predict relative chemical reactivities of congeneric compounds and regio- and stereoselectivities in chemical reactions.[1] Canonical molecular orbitals (CMOs), which are used in the FMO theory, are delocalized in molecules. Therefore, it is necessary to consider not only energy levels but densities of FMOs at the reaction site in order to predict relative chemical reactivities of congeneric molecules. Moreover, it is also necessary to consider CMOs other than FMO when the energy levels of the CMOs are close to that of the FMO and the densities of the CMOs at the reaction site are not negligible. In the FMO theory, superdelocalizabilities are used to overcome these problems.[2] However, the superdelocalizability is defined only at an atom, and the phase of CMOs cannot be considered. Therefore, multicentered reactions cannot be analyzed with the superdelocalizabilities.

Fujimoto et al. defined the localized frontier orbital (LFO).[3] Since the LFO is localized at the reaction site, only energy levels of LFOs are considered to predict relative chemical reactivities of congeneric molecules. The effects of all CMOs that have densities at the reaction site are included in the LFO energies. Multicentered reactions can be analyzed using LFOs if appropriate reference orbitals are chosen.

An LFO can be obtained by projecting a certain reference orbital onto the occupied or unoccupied molecular orbital space of a molecule.[3] When the density of a π orbital at an atom is important for a reaction and a minimal basis set is used for the molecular orbital calculation, the reference orbital is simply the valence p_π atomic orbital of the atom. However, there was an ambiguity in the selection of the reference orbital when lone pair orbitals or σ^* orbitals were dealt with. There was also an ambiguity when a split valence basis set was used. Although Omoto et al. minimized the energy of the unoccupied LFO to find an appropriate combination of the inner and outer parts of Si 3p functions for the reference orbital at the HF/6-31G** level, they did not show the method for the minimization.[4] Perhaps they did not specify the method because that was a numerical one.

In this paper, we adopted the variational method to analytically find appropriate combinations of valence atomic orbitals giving the maximum and minimum energies of the occupied and unoccupied LFOs, respectively. We have used this method to investigate correlations between LFO energies and acidities of conjugate cations of heterocyclic bases and monosubstituted benzoic acids.

Method

Variational Equation. The reference orbital δ_r is expanded in terms of the atomic orbitals in the reaction center.

$$\delta_r = \sum_\mu C_\mu \chi_\mu \tag{1}$$

χ_μ is the μth atomic orbital in the reaction center, and C_μ is the expansion coefficient for χ_μ. It is not necessary that all χ_μ's ($\mu = 1, ...$) belong to the same atom. In the case that δ_r is a σ^* orbital of an OH bond, χ_μ's ($\mu = 1, ...$) are valence atomic orbitals of the O and H. By using eq 1, we can also deal with the sp hybridization for a lone pair and the combination of inner and outer valence atomic orbitals for split valence basis sets. The atomic orbital χ_μ can be expanded in terms of the occupied canonical molecular orbitals ϕ_i and the unoccupied canonical molecular orbitals ϕ_j in the LCAO approximation.

$$\chi_\mu = \sum_i^{oc} D_\mu^i \phi_i + \sum_j^{unoc} D_\mu^j \phi_j \tag{2}$$

Equation 3 can be derived from eqs 1 and 2.

$$\delta_r = \sum_i^{oc} d_{ir} \phi_i + \sum_j^{unoc} d_{jr} \phi_j \tag{3}$$

$$d_{ir} = \sum_\mu C_\mu D_\mu^i$$

$$d_{jr} = \sum_\mu C_\mu D_\mu^j$$

Energies of occupied and unoccupied LFOs are defined by eqs 4 and 5, respectively, where ϵ_i is the energy of the ith canonical molecular orbital.[3]

$$\lambda_{oc}(\delta_r) = (\sum_i^{oc} d_{ir}^2 \epsilon_i)/(\sum_i^{oc} d_{ir}^2) \tag{4}$$

$$\lambda_{unoc}(\delta_r) = (\sum_j^{unoc} d_{jr}^2 \epsilon_j)/(\sum_j^{unoc} d_{jr}^2) \tag{5}$$

To analytically obtain a set of C_μ's ($\mu = 1, ...$) that gives the

† E-mail: kurita@tziris3.sumitomo-chem.co.jp.
‡ E-mail: takayama@tziris3.sumitomo-chem.co.jp.
⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1997.

S1089-5639(97)00922-5 CCC: $14.00

maximum value of $\lambda_{oc}(\delta_r)$ or the minimum value of $\lambda_{unoc}(\delta_r)$, we have derived eq 6 from eqs 3, 4, and 5 using the variational method.

$$\mathbf{EC} = \mathbf{SC}\boldsymbol{\lambda} \tag{6}$$

$$(\mathbf{E})_{\mu\nu} = \sum_{i}^{oc} D_{\mu}^{i} D_{\nu}^{i} \epsilon_i \quad \text{for} \quad \lambda_{oc}(\delta_r)$$

$$\sum_{j}^{unoc} D_{\mu}^{j} D_{\nu}^{j} \epsilon_j \quad \text{for} \quad \lambda_{unoc}(\delta_r)$$

$$(\mathbf{S})_{\mu\nu} = \sum_{i}^{oc} D_{\mu}^{i} D_{\nu}^{i} \quad \text{for} \quad \lambda_{oc}(\delta_r)$$

$$\sum_{j}^{unoc} D_{\mu}^{j} D_{\nu}^{j} \quad \text{for} \quad \lambda_{unoc}(\delta_r)$$

$\boldsymbol{\lambda}$ is a diagonal matrix and $(\boldsymbol{\lambda})_{nn}$ is the energy of the nth LFO in the reaction center. $(\mathbf{C})_{\mu n}$ is the expansion coefficient of the nth reference orbital for χ_μ.

Molecular Orbital Calculations. Geometries of the heterocyclic bases used in this work were optimized at the HF/6-31G* level. Conformational analyses were carried out for monosubstituted benzoic acids at the HF/6-31G*//HF/3-21G level, and the geometries of the most stable conformer of each compound were further optimized at the HF/6-31G* level. Molecular orbitals at the HF/STO-3G, HF/6-31G, HF/6-31G*, and HF/6-31G** levels were obtained by using these HF/6-31G* level geometries. The SPARTAN[5] and Gaussian 94[6] programs were used for these calculations.

Results and Discussion

Acidities of Conjugate Cations of Heterocyclic Bases. Basicities of sp^2-type nitrogens in heterocyclic bases are governed by the characteristics of the lone pair orbitals of the nitrogens. The highest occupied LFO obtained by eq 6 at an sp^2-type nitrogen in heterocyclic bases is a π-type orbital. Therefore, we selected the second highest occupied LFO at the nitrogen, which was the lone pair orbital, to estimate the basicity of the nitrogen. We designate the energy of the second highest occupied LFO by $\lambda_{oc}(\delta_r)_{lp}$. A combination of all s- and p-type valence atomic orbitals at the basic nitrogen was considered to maximize $\lambda_{oc}(\delta_r)_{lp}$. The d-type atomic orbitals were not considered at the HF/6-31G* level because $\mathbf{S}^{-1/2}$, which was necessary to solve eq 6, could not be obtained when the d-type AOs were included.

Measured pK_a values of conjugate cations of heterocyclic bases taken from the literature[7] and the maximum value of $\lambda_{oc}(\delta_r)_{lp}$'s at sp^2-type nitrogens in each base ($\lambda_{oc}(\delta_r)_{lp}$·max) are listed in Table 1. Simple correlation coefficients between measured pK_a values and $\lambda_{oc}(\delta_r)_{lp}$·max are shown in Table 2. Standard deviations of pK_a values predicted with the correlation equations from measured pK_a values are also shown in Table 2. At the HF/STO-3G level, the simple correlation coefficient is much lower than those at the HF/6-31G and HF/6-31G* levels. The correlation at the HF/6-31G level is slightly better than that at the HF/6-31G* level.

At each level, the correlation becomes better when 1,2,3-triazole is considered for the correlation analysis instead of 1,2,5-triazole. At the HF/6-31G* level, 1,2,5-triazole is 4.9 kcal/mol more stable than 1,2,3-triazole. However, the stabilization energy of 1,2,3-triazole by protonation is 18.4 kcal/mol larger than that of 1,2,5-triazole at the HF/6-31G* level. This is consistent with the better correlation considering 1,2,3-triazole.

In the correlation considering 1,2,3-triazole at the HF/6-31G level (Figure 1), the difference between measured and predicted pK_a values is large for 2,1-benzisothiazole (−2.47), isothiazole (−2.32), and 1,2,5-thiadiazole (−1.68). At each level, the correlation becomes better when these compounds in which sulfur atoms are adjacent to sp^2-nitrogen atoms are excluded (Table 2). The reason for this is still not clear.

Acidities of Monosubstituted Benzoic Acids. Acidities of benzoic acids are governed by the characteristics of the OH σ^* orbitals in carboxyl groups. The second lowest unoccupied LFO

TABLE 1: Measured pK_a Values of Conjugate Cations of Heterocyclic Bases and the Maximum Value of $\lambda_{oc}(\delta_r)_{lp}$'s at sp^2-Type Nitrogens in Each Base ($\lambda_{oc}(\delta_r)_{lp}$·max)

		$\lambda_{oc}(\delta_r)_{lp}$·max (eV)		
molecule	pK_a [a]	HF/STO-3G	HF/6-31G	HF/6-31G*
1,2,5-thiadiazole	−4.9	−14.429	−14.726	−14.571
1,2-benzisoxazole	−4.7	−15.574	−14.859	−14.800
isoxazole	−2.97	−15.599	−14.820	−14.700
2,1-benzisoxazole	−2.20	−15.320	−14.608	−14.526
isothiazole	−0.51	−13.487	−13.838	−13.818
benzoxazole	−0.13	−14.368	−14.167	−14.071
2,1-benzisothiazole	−0.05	−13.393	−13.731	−13.727
pyrazine	0.4	−13.993	−13.966	−13.905
1-methylindazole	0.42	−14.690	−14.054	−14.070
2-chloropyridine	0.7	−14.047	−13.899	−13.846
oxazole	0.8	−14.325	−14.040	−13.877
pyrimidine	1.1	−13.905	−13.878	−13.764
1,2,3-triazole (N_3)[b]	1.17	−14.595	−14.117	−14.027
(1,2,5-triazole)	(1.17)	(−15.768)	(−15.021)	(−14.891)
benzothiazole	1.2	−13.723	−13.932	−13.865
1-methyl-1,2,3-triazole (N_3)[b]	1.25	−14.425	−13.918	−13.862
indazole	1.31	−14.868	−14.203	−14.207
2-methylindazole	2.02	−14.349	−13.794	−13.826
1-methylpyrazole	2.06	−14.452	−13.792	−13.798
pyridazine	2.1	−14.093	−13.957	−13.866
1,2,4-triazole (N_4)[b]	2.45	−14.342	−13.907	−13.805
pyrazole	2.52	−14.675	−13.976	−13.963
thiazole	2.53	−13.615	−13.831	−13.733
3-chloropyridine	2.8	−13.751	−13.666	−13.645
1-methyl-1,2,4-triazole (N_4)[b]	3.20	−14.189	−13.714	−13.651
4-chloropyridine	3.8	−13.747	−13.637	−13.604
pyridine	5.2	−13.175	−13.180	−13.195
benzimidazole	5.53	−13.548	−13.295	−13.342
1-methylbenzimidazole	5.57	−13.487	−13.183	−13.257
3-methylpyridine	5.7	−13.084	−13.102	−13.127
2-methylpyridine	6.0	−12.965	−13.014	−13.023
4-methylpyridine	6.0	−13.100	−13.066	−13.100
imidazole	6.95	−13.337	−13.021	−13.043
1-methylimidazole	7.33	−13.251	−12.889	−12.945

[a] Measured pK_a values are taken from ref 7. [b] The basic nitrogen at which $\lambda_{oc}(\delta_r)_{lp}$ is maximum.

TABLE 2: Simple Correlation Coefficients (r) between Measured pK_a Values and $\lambda_{oc}(\delta_r)_{lp}$·max and Standard Deviations (s) of pK_a Values Predicted with the Correlation Equations from Measured pK_a Values for Heterocyclic Bases

	HF/STO-3G	HF/6-31G	HF/6-31G*
r^a	0.76	0.95	0.95
s^a	2.02	0.95	1.00
r^b	0.72	0.90	0.89
s^b	2.17	1.37	1.40
r^c	0.89	0.98	0.97
s^c	1.34	0.65	0.72

[a] 1,2,3-Triazole was considered for the correlation analysis instead of 1,2,5-triazole. [b] 1,2,5-Triazole was considered for the correlation analysis instead of 1,2,3-triazole. [c] 1,2,3-Triazole was considered for the correlation analysis instead of 1,2,5-triazole. 1,2,5-Thiadiazole, isothiazole, and 2,1-benzisothiazole were not included in the correlation analysis.

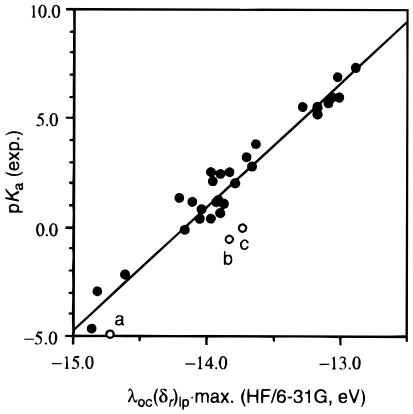

Figure 1. Correlation between measured pK_a values and $\lambda_{oc}(\delta_r)_{lp}$·max at the HF/6-31G level for heterocyclic bases. 1,2,3-Triazole is considered for the correlation instead of 1,2,5-triazole. Open circles a, b, and c show the values for 1,2,5-thiadiazole, isothiazole, and 2,1-benzisothiazole, respectively.

TABLE 3: Hammett σ Constants of Substituents and $\lambda_{unoc}(\delta_r)_{\sigma^*}$ for Monosubstituted Benzoic Acids

		$\lambda_{unoc}(\delta_r)\sigma^*$ (eV)			
substituent	σ_m or σ_p[a]	HF/ STO-3G	HF/ 6-31G	HF/ 6-31G*	HF/ 6-31G**
p-1-pyrrolidinyl	−0.90	16.480	6.579	6.810	6.826
p-$N(CH_3)_2$	−0.83	16.415	6.528	6.759	6.775
p-$NHCH_3$	−0.70	16.427	6.537	6.776	6.790
p-NH_2	−0.66	16.410	6.512	6.758	6.771
p-OH	−0.37	16.149	6.279	6.546	6.562
p-OCH_3	−0.27	16.193	6.337	6.591	6.606
m-$NHCH_3$	−0.21	16.108	6.325	6.566	6.581
p-CH_3	−0.17	16.092	6.315	6.544	6.560
m-NH_2	−0.16	16.128	6.339	6.582	6.597
m-$N(CH_3)_2$	−0.16	16.153	6.369	6.602	6.617
m-CH_3	−0.07	16.049	6.283	6.513	6.529
H	0.00	16.016	6.249	6.484	6.498
p-SCH_3	0.00	16.122	6.238	6.483	6.498
p-F	0.06	15.949	6.087	6.368	6.382
m-OH	0.12	15.927	6.139	6.406	6.422
m-OCH_3	0.12	15.980	6.203	6.460	6.475
m-SCH_3	0.15	15.978	6.149	6.397	6.411
p-Cl	0.23	15.693	6.040	6.296	6.310
m-F	0.34	15.848	6.049	6.318	6.332
p-OCF_3	0.35	15.870	5.997	6.309	6.324
m-$COOCH_3$	0.36	15.909	6.128	6.356	6.371
m-Cl	0.37	15.666	6.030	6.283	6.296
m-$COCH_3$	0.38	15.867	6.034	6.281	6.296
m-OCF_3	0.38	15.750	5.957	6.260	6.274
m-CF_3	0.43	15.749	5.944	6.239	6.252
p-$COOCH_3$	0.45	15.857	6.087	6.326	6.341
p-$COCH_3$	0.50	15.869	6.055	6.298	6.314
p-CF_3	0.54	15.721	5.901	6.206	6.221
m-CN	0.56	15.587	5.886	6.113	6.126
p-CN	0.66	15.551	5.856	6.078	6.092
m-NO_2	0.71	15.516	5.816	6.083	6.096
p-NO_2	0.78	15.442	5.731	6.002	6.017
m-SO_2CF_3	0.83	15.468	5.683	6.024	6.037
p-SO_2CF_3	0.96	15.391	5.568	5.929	5.943
r[b]		0.966	0.965	0.975	0.975
s[c]		0.12	0.12	0.10	0.10

[a] Hammett σ constants are taken from ref 8. [b] Simple correlation coefficients between Hammett σ constants and $\lambda_{unoc}(\delta_r)_{\sigma^*}$. [c] Standard deviations of Hammett σ constants predicted with the correlation equations from measured Hammett σ constants.

obtained by eq 6 at the carboxyl OH group is the OH σ^* orbital at the HF/STO-3G and HF/6-31G levels. On the other hand, the lowest unoccupied LFO obtained by eq 6 at the carboxyl OH group is the OH σ^* orbital at the HF/6-31G* and HF/6-31G** levels. We designate the energy of the OH σ^* orbital by $\lambda_{unoc}(\delta_r)_{\sigma^*}$. A combination of all s-, p-, and d-type valence atomic orbitals of the O and H in the carboxyl OH group was considered to minimize $\lambda_{unoc}(\delta_r)_{\sigma^*}$. We represent the acidities of benzoic acids by Hammett σ constants defined as the differences in pK_a values between benzoic acid and substituted benzoic acids. Hammett σ constants[8] and $\lambda_{unoc}(\delta_r)_{\sigma^*}$ are listed in Table 3. The basis set effect on the correlation between Hammett σ constants and $\lambda_{unoc}(\delta_r)_{\sigma^*}$ is smaller than that on the correlation between pK_a values and $\lambda_{oc}(\delta_r)_{lp}$·max. The correlation at the HF/6-31G* or HF/6-31G** level is slightly better than that at the HF/STO-3G or HF/6-31G level. The correlation at the HF/6-31G** level is shown in Figure 2.

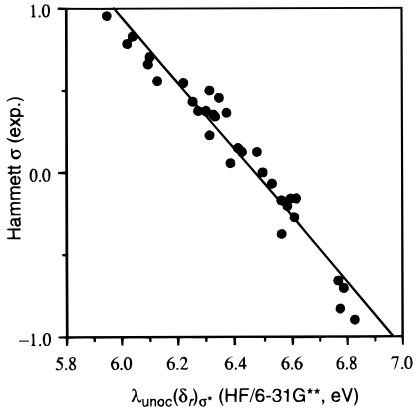

Figure 2. Correlation between Hammett σ constants and $\lambda_{unoc}(\delta_r)_{\sigma^*}$ at the HF/6-31G** level for monosubstituted benzoic acids.

Conclusion

The variational method has been applied to the analytical determination of reference orbitals for localized frontier orbitals. An equation similar to the Roothaan−Hartree−Fock equation has been obtained. This equation was used to obtain lone pair orbital energies for sp^2-type nitrogen atoms in heterocyclic bases and carboxyl OH σ^* orbital energies for monosubstituted benzoic acids. Good linear correlations have been found between localized frontier orbital energies and acidities of conjugate cations of heterocyclic bases and monosubstituted benzoic acids.

References and Notes

(1) Fukui, K. *Theory of Orientation and Stereoselection*; Springer-Verlag: Berlin, 1975.

(2) Fukui, K.; Yonezawa, T.; Nagata, C. *Bull. Chem. Soc. Jpn.* **1954**, *27*, 423.

(3) Fujimoto, H.; Mizutani, Y.; Iwase, K. *J. Phys. Chem.* **1986**, *90*, 2768.

(4) Omoto, K; Sawada, Y.; Fujimoto, H. *J. Am. Chem. Soc.* **1996**, *118*, 1750.

(5) SPARTAN (Version 3.0), Wavefunction, Inc., Irvine CA, 1993.

(6) Gaussian 94 (Revision A.1), Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1995.

(7) Katritzky, A. R. *Handbook of Heterocyclic Chemistry*; Pergamon Press: New York, 1985.

(8) Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR (Hydrophobic, Electronic, and Steric Constants)*; American Chemical Society: Washington, DC, 1995.